The C-terminal Domains of Two Homologous Oleaceae β-1,3-Glucanases Recognize Carbohydrates Differently: Laminarin Binding by NMR

Héctor Zamora-Carreras[1], María Torres[2], Noemí Bustamante[1], Anjos L. Macedo[3], Rosalía Rodríguez[2], Mayte Villalba[2], and Marta Bruix[1]

[1] Departamento de Química Física Biológica, Instituto de Química Física "Rocasolano", CSIC, Serrano 119, 28006 Madrid, Spain.

[2] Departamento de Bioquímica y Biología Molecular I, Facultad de Química, Universidad Complutense, 28040 Madrid, Spain.

[3] UCIBIO, REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal,

Corresponding author: Prof. M. Bruix, Departamento de Química Física Biológica, Instituto de Química Física "Rocasolano", Serrano 119, 28006 Madrid, Spain. Phone: +34 91 745 9511. Fax: +34 91 564 24 31, E-mail: mbruix@iqfr.csic.es

Running title: CtD-Ole e 9 and CtD-Fra e 9-laminarin recognition

Keywords: Fra e 9, Ole e 9, allergy, ash pollen, β-1,3-glucanase, carbohydrate-binding protein, NMR

ABSTRACT

Ole e 9 and Fra e 9 are two allergenic β-1,3-glucanases from olive and ash tree pollens, respectively. Both proteins present a modular structure with a catalytic N-terminal domain and a carbohydrate-binding module (CBM) at the C-terminus. Despite their significant sequence resemblance, they differ in some functional properties, such as their catalytic activity and the carbohydrate-binding ability. Here, we have studied the different capability of the recombinant C-terminal domain of both allergens to bind laminarin by NMR titrations, binding assays and ultracentrifugation. We show that rCtD-Ole e 9 has a higher affinity for laminarin than rCtD-Fra e 9. The complexes have different exchange regimes on the NMR time scale in agreement with the different affinity for laminarin observed in the biochemical experiments. Utilizing NMR chemical shift perturbation data, we show that only one side of the protein surface is affected by the interaction and that the binding site is located in the inter-helical region between α1 and α2, which is buttressed by aromatic side chains. The binding surface is larger in rCtD-Ole e 9 which may account for its higher affinity for laminarin relative to rCtD-Fra e 9.

INTRODUCTION

Endo-β-1,3-glucanases are a family of glucan hydrolases (EC 3.2.1.39) widely distributed among plants, fungi and bacteria. They catalyze the hydrolytic cleavage of 1,3-D-glucosidic linkages in β-1,3-glucans and, on a larger scale, play a determining role in structural molecular conversions or metabolic functions such as remodelling cell walls, cell expansion processes, cell-cell fusion, as well as cell division (Verma and Hong 2005), microsporogenesis (Worrall, Hird et al. 1992), pollen tube growth (Vogeli-Lange, Frundt et al. 1994) or seed germination (Casacuberta, Raventos et al. 1992). β-1,3-glucanases from higher plants belong to the glycosylhydrolase family 17 (GHF 17), whose components differ in molecular properties, cellular location and expression pattern. In addition to their constitutive biochemical functions, a defensive role has been attributed to these enzymes, as they can be expressed as a response to pathogens (Linthorst, Melchers et al. 1990, Memelink, Linthorst et al. 1990). Based on these findings, they have been classified as the pathogen-related group 12 (PR-12). The antifungal activity of β-1,3-glucanases from higher plants has encouraged scientists to consider using molecular bioengineering to modify these enzymes with the goal of developing fungi-resistant crops (Chen and Chen 2000).

Concerning their molecular structure, two main types of β-1,3-glucanases have been reported in plants: i) long enzymes with 42-50 kDa molecular masses, composed by two domains -a large N-terminal domain (NtD, 33-40 kDa) with catalytic activity and a small C-terminal domain (CtD, around 10 kDa) with carbohydrate-binding capacity which has been called carbohydrate-binding module (CBM); and ii) short β-1,3-glucanases (33-41 kDa) in which the CtD is lacking. The two domains from long β-1,3-glucanases fold independently and can be expressed recombinantly in heterologous systems. The three-dimensional modelling of all known NtD of β-1,3-glucanases closely resemble the canonical triose-phosphate isomerase (TIM)-barrel structure, although proteins from phylogenetically non-related species do not show high sequence identity.

In addition to the biochemical activity of β-1,3-glucanases, allergenic properties have been described for several members of this protein family. β-1,3-glucanases from pollens, fruits and natural latex (*Hevea brasiliensis)* are able to trigger allergic symptoms in hypersensitive patients. Ole e 9 and Fra e 9 are allergenic β-1,3-glucanases

from olive tree (*Olea europaea*) and ash (*Fraxinus excelsior*) pollens, respectively; both species belonging to the *Oleaceae* family. The N- and C-terminal domains from Ole e 9 have been molecular and immunologically characterized (Huecas, Villalba et al. 2001). Also, the solution structure of the CtD-Ole e 9, a domain belonging to the CBM43 family, has been determined (Treviño, Palomares et al. 2008). Moreover, we have recently reported the cloning, sequencing and the independently recombinant expression of the two domains of Fra e 9 which are composed of 320 residues for the NtD and 108 residues for the CtD (Torres 2014). The identity between rCtD-Fra e 9 and rCtD-Ole e 9 sequences calculated by the SIM alingment tool (Huang and Miller 1991) is 55.7%, (Figure 1A).

It has been postulated that CtD acts as to capture the β-1,3-glucan substrate which is then hydrolyzed by the catalytic NtDs (Boraston, Bolam et al. 2004). However, a detailed description of the process is still unknown. We expect that a high-resolution study of protein-ligand interactions would lead to a better understanding of the molecular basis of the biological recognition and functional mechanisms.

In this manuscript, we report the structural, hydrodynamic characterisation and the carbohydrate binding capability of rCtD-Fra e 9 from ash pollen in comparison with those of its olive tree pollen counterpart, Ole e 9, in order to understand its role in the catalytic process of long β-1,3-glucanases.

MATERIALS and METHODS

Materials.

All ligands used for the binding assays were purchased from Sigma-Aldrich (USA): agarose, CM-cellulose (carboxymethil cellulose), laminaritetraose (purity ≥ 90%), laminarihexaose (purity ≥ 99%), laminarin (from *Laminaria digitata*) and lichenan (from *Cetraria islandica*).

Protein production and purification.

rCtD-Fra e 9, which comprises residues D354-S461 of Fra e 9, was produced in *Pichia pastoris* strain KM71 as previously described (Palomares, Villalba et al. 2003). Briefly, cells were grown in 1 L of BMG (100 mM K_2HPO_4 pH 6, 0.34% yeast nitrogen base, 1% $(NH_4)_2SO_4$, $4 \cdot 10^{-5}$% biotin and 1% glycerol) for 72 h at 30 ºC . Then, cells were grown in 200 mL of induction medium BMM (100 mM K_2HPO_4, pH 6, 0.34% yeast nitrogen base, 1% $(NH_4)_2SO_4$, $4 \cdot 10^{-5}$% biotin and 0.5% methanol). After 4 days,

the supernatant was dialyzed in the presence of 20 mM NH_4HCO_3. Two chromatography steps: (i) gel filtration Sephadex G-50 column in 0.2 M NH_4HCO_3 and (ii) Nucleosil C-18 (RP-HPLC) TFA 0.1% with a gradient of acetonitrile (0-60%), were used for the protein purification. The purity was analysed by 15% SDS-PAGE. rCtD-Ole e 9 was produced and purified as previously described (Treviño, Palomares et al. 2008).

To produce ^{15}N-^{13}C labelled proteins, the same procedure was employed with the slight modifications described previously by (Treviño, García-Mayoral et al. 2004); namely $(NH_4)_2SO_4$ was substituted by $(^{15}NH_4)_2SO_4$ (Cambridge Isotopes) in the BMG and BMM media, and 0.5% ^{13}C-glucose (Cambridge Isotopes) was used instead of glycerol in BMG and methanol was replaced by ^{13}C-methanol (Cambridge Isotopes) in BMM. All the samples were analysed by amino acid analysis, N-terminal sequencing and mass spectroscopy.

Carbohydrate-binding assay.

The polysaccharide-binding activity of the rCtD from Fra e 9 and Ole e 9 was tested by affinity gel electrophoresis (AGE) as described previously (Barral, Suarez et al. 2005). Proteins (2 μg) were electrophoresed in native 15% polyacrylamide gels containing laminarin or lichenan ranging from 0.062 to 1.2 mM or 0.24 to 4.8 mM, respectively. Other different soluble oligosaccharides and polysaccharides with β(1→3) (laminaritetraose and laminarihexaose) and β(1→4) (agarose and CM-cellulose) linkages were also soaked up in the separating gel mixture at a concentration of 2.5 $mg \cdot mL^{-1}$ prior to polymerization. Gels without ligand were electrophoresed simultaneously. BSA (0.7 μg) was used as a negative control. The K_D value for the binding of rCtD-Fra e 9 to ligand under the conditions described was determined as the inverse of the absolute value of the intersection of the plot with the abscissa (Bolam, Xie et al. 2004).

NMR spectroscopy and spectral assignment.

rCtD-Ole e 9 and rCtD-Frae e 9 samples were prepared at 0.4 mM in H_2O/D_2O (9:1 v/v) containing sodium-4,4-dimethyl-4-silapentane-1-sulfonate (DSS) as the internal ^1H chemical shift reference at pH 6.0. NMR spectra were acquired at 298 K on a Bruker AV 800 NMR spectrometer equipped with a triple-resonance cryoprobe and an active shielded z-gradient coil. ^{15}N and ^{13}C chemical shifts were referenced indirectly

using the gyromagnetic ratios of ^{15}N:^{1}H and^{13}C:^{1}H (Wishart, Bigam et al. 1995). Standard 2D ^{15}N-HSQC and ^{13}C-HSQC and 3D spectra CBCA(CO)NH, CBCANH, HNCA, HN(CO)CA were acquired and analysed.

The spectra were processed with Bruker Topspin (Bruker, Germany) and spectral analysis was performed with Sparky3 (Goddard and Kneller 2005). Backbone assignment of rCtD-Fra e 9 was performed following conventional strategies as previously described for rCtD-Ole e 9 (Treviño, Palomares et al. 2008).

Titration of olive tree and ash rCtDs with laminarin monitored by NMR.

Increasing amounts of a laminarin (mean molecular mass about 5.5 kDa) solution from *Laminaria digitata* (20 mM and pH 6.0) were added to ^{13}C-^{15}N rCtD-Ole e 9 and ^{13}C-^{15}N rCtD-Fra e 9 sample solutions (0.4 mM and pH 6.0) and series of ^{15}N-HSQC spectra were recorded at each titration point at 298 K. The final titration point was set at a concentration of ≈1mM of laminarin because precipitation was observed at higher values. Changes of peak intensity and position were monitored, and both the chemical shift and line width changes were analysed. In all cases, the pH was measured at the final points of the titrations. For the mapping of the interaction surface, average amide ^{15}N and ^{1}H chemical shift perturbations were calculated according to the equation: $\delta_{av}= (((\Delta\delta_{NH})^2 + (\Delta\delta_N)^2/25))/2)^{0.5}$.

NMR dynamics.

NMR relaxation experiments were carried out at the same conditions described above for the laminarin-free and bound proteins. Conventional ^{15}N heteronuclear relaxation rates T_1, T_2 and NOE data were determined. To this end, a series of 2D heteronuclear correlated spectra using sensitivity enhanced gradient pulse scheme (Farrow, Muhandiram et al. 1994) were recorded. The relaxation delay times were set as follows for T_1: 5, 50, 150, 300, 600, 800, 1000 and 1200 ms; and for T_2: 16, 32, 48, 64, 80, 96, 112 and 128 ms. The relaxation time constants T_1 and T_2 were obtained from the exponential fits of the measured crosspeak intensities. The uncertainty was taken as the error in the fit of the decay function. For the NOE measurement, the experiments with and without proton saturation were acquired simultaneously in an interleaved manner with a recycling delay of 5 s, and were split during processing into separate spectra for analysis. The values for the heteronuclear NOEs were obtained from the ratio intensities

of the resonances according to: I_{sat}/I_{ref}. Here, the uncertainty was estimated to be about 5%.

The correlation times (τ_c) were estimated for both rCtD-Fra e 9 rCtD-Ole e 9 (free and the laminarin complexes) from experimental $<T_1/T_2>$ values using the program HydroNMR (García de la Torre, Huertas et al. 2000). According to the literature (Kay, Torchia et al. 1989), the experimental $<T_1/T_2>$ ratios were modified by excluding those residues with T_1, T_2 values that deviate more than one standard deviation from the mean, and with NOE values below 0.65. Calculations with HydroNMR were performed assuming a globular shape and a rigid behaviour for both the isolated domains and complexes. Theoretical correlation times were estimated by a back-calculation procedure based on an iterative method that allows us to compare the theoretical $<T_1/T_2>$ values with the experimental ones obtained as described above.

The molecular masses of the rCtD-Ole e 9 and rCtD-Fra e 9-laminarin complexes were estimated from the corresponding correlation time values through an interpolation based on the least-squares fit of a linear function to experimental correlation times obtained for different proteins (Aramini 2010).

Analytical ultracentrifugation.

Sedimentation and velocity experiments were carried out using an Optima XL-A analytical ultracentrifuge (absorption optics) at 25ºC. Samples were prepared in the same conditions used for NMR experiments, H_2O/D_2O (9:1 v/v) pH 6.0, in the presence of 0.91 mM of laminarin (aprox. 5.5 kDa molecular mass) and using concentrations of $3.16 \cdot 10^{-5}$ M and $3.77 \cdot 10^{-5}$ M for rCtD-Ole e 9 (10,509.7 Da) and rCtD-Fra e 9 (11,364.4 Da), respectively.

Equilibrium assays were performed by centrifugation of 80 µL of each sample at 19,000, 22,400 and 33,000 rpm, checking mass conservation for each velocity. Sedimentation profiles were analysed following a single species sedimentation model as previously described (Varea, Saiz et al. 2000). The SEDNTERP program (Laue, Shah et al. 1992) was used to calculate the protein specific volume from the sequences (0.7095 $cm^3 \cdot g^{-1}$ for rCtD-Ole e 9 and 0.7076 $cm^3 \cdot g^{-1}$ for rCtD-Fra e 9) as well as the buffer viscosity and density. For velocity measurements (45,000 rpm) 400 µL of samples were used. Differential sedimentation coefficients, c(s), were calculated by least squares boundary modelling of sedimentation velocity profiles using the program SEDFIT

(Schuck 2000) and standard sedimentation coefficients ($S_{20,w}$) were calculated by SEDNTERP from experimental values.

rCtD-Fra e 9 model.

Starting from the reported structure of rCtD-Ole e 9 (PDB code: 2JON) (Treviño, Palomares et al. 2008), and on the basis of the significant sequence identity and similar NMR parameters between both proteins, a structural model of rCtD-Fra e 9 was constructed using the SWISS-MODEL server (Guex, Peitsch et al. 2009). To perform an accurate modelling, the unstructured N-terminal tails of rCtD-Ole e 9 (residues A360-S371) and rCtD-Fra e 9 (resiudes D354-K374) were not considered in this process.

RESULTS

NMR spectral assignment, secondary structure and global fold of rCtD-Fra e 9.

The spectra of the rCtD-Fra e 9 were assigned at pH 6, following the standard NMR hetreonuclear methodology. Assignment of the ^{13}C and ^{15}N backbone chemical shifts was facilitated by comparing rCtD-Fra e 9 data with the reported assignment of the homologous rCtD-Ole e 9 (Castrillo, Treviño et al. 2006), on the basis of the sequence similarity (**Figure 1A**). The assignment is nearly complete (**Figure 1B**), with the exception of eight residues (in black) near the N-terminus corresponding to a Pro rich region: ^{356}**PV**PT**PSS**PV**PKP**^{367}. To evaluate the secondary structure, C_α and C_β conformational shifts ($\Delta\delta$) were calculated for every nuclei as the difference between the chemical shift observed experimentally and a reference value obtained for random coil peptides (Wishart, Bigam et al. 1995). Significant positive C_α conformational shifts indicative of helical structure (Wishart, Sykes et al. 1991) are found in the stretches D384 to S396, V417 to S431 and D436 to G438; whereas significant negative C_α conformational shifts characteristic of β-strand structure were found in C376 to P378 and G445 to T448. These results obtained for the rCtD-Fra e 9 were compared with those from the rCtD-Ole e 9 and reveal the high similarity in the number, type and position of the elements of secondary structure (**Figure 1B**). Overall, chemical shift differences are small which strongly suggests that the secondary structure and the global fold of rCtD-Ole e 9 is preserved in the rCtD-Fra e 9.

Carbohydrate binding activity of rCtD-Fra e 9.

The ability of purified rCtD-Fra e 9 to bind soluble polysaccharides was assessed by quantitative AGE under non-denaturing conditions. Carbohydrates of different length or linkages (β-1,3 or β-1,3/1,4) -agarose, CM-cellulose, lichenan, laminaritetraose, laminarihexaose and laminarin- were assayed. Relative to rCtD-Ole e 9, purified rCtD-Fra e 9 displayed a significant but lower specific binding to laminarin (**Figure 2A**) as demonstrated by the clear shift of mobility for this domain in AGE. The rCtD-Fra e 9 did not show significant binding capacity to any of the other polysaccharides assayed, except a weak mobility change in the presence of lichenan (β-1,3/1,4-glucan) (**Figure 2B**). The reciprocal relative migration distance (1/(R-r)) was plotted against the carbohydrate concentration (**Figure 2C and 2D**). rCtD-Fra e 9-laminarin complex shows a K_D of 1.1±0.4 mM, indicating that its interaction is considerably weaker than that of rCtD-Ole e 9-laminarin (K_D = 0.065 ±0.003 mM).

Testing the interaction of rCtD-Fra e 9 and rCtD-Ole e 9 with laminarin by NMR.

The interaction of rCtD-Fra e 9 and rCtD-Ole e 9 with laminarin was first tested comparing the ^{15}N-HSQC spectra recorded for free proteins and for the (~1:1 and 1:2.5) complexes in samples where only the protein moiety was labelled. As explained before, we could not use a higher excess of laminarin due to precipitation problems. This means that, in the case of rCtD-Fra e 9, we did not reach a concentration of laminarin above the calculated K_D value (≈1mM < 1.1 mM).

For rCtD-Fra e 9, a significant number of resonances shift, as seen in **Figure 3A,** confirming the direct interaction in a fast exchange regime. To facilitate detection of the residues most affected by complex formation, we used the weighted average of amide ^{15}N and ^{1}H chemical shift variations (**Figure 4A**), and mapped the protein interacting surface (**Figure 4B, 4C**) into the protein's structural model. The rCtD-Fra e 9 residues most affected by the interaction ($\Delta\delta_{avg}$ > mean value) are: K373, S389, I391-D392, V394-S396, G398-G399, V405, A407, N415, A419-A421, Y423-M425, W428-N435, G438, F441, G445 and S450.

Following the addition of laminarin to rCtD-Ole e 9, and upon complex formation, a set of ^{15}N-^{1}H NMR signals corresponding to the interacting region shift, for instance: W372, G386, N389-C392, I396, V420-M421, G429, D436-A441; and some others become broad and completely disappear from the spectra: G404, H427-A428, S432, N434 and C435. This fact corroborates the direct interaction of rCtD-Ole e 9 with laminarin on the medium-slow exchange regime in the NMR time scale

(milliseconds or slower). Similarly to what was found for rCtD-Fra e 9, in rCtD-Ole e 9 chemical shift perturbations map the laminarin-binding site to helices α1, α2 and α3. In addition, the following loop and the strand β2 of rCtD-Ole e 9 are also affected by the complex formation (**Figure 4A, 4C**). Remarkably the resonances that suffer severe broading effects are next to the long and unstructured N-terminal tail in the rCtD-Ole e 9 structure.

The interaction between allergen domains and laminarin was also tested using the NMR relaxation data obtained for both rCtD-Fra e 9 and rCtD-Ole e 9. As expected, globally T_1 values increases and T_2 values decreases as a consequence of molecular association. For rCtD-Fra e 9, mean values are $\langle T_1 \rangle = 0.64$ s and $\langle T_2 \rangle = 0.097$ s for the free protein, and $\langle T_1 \rangle = 0.96$ s and $\langle T_2 \rangle = 0.051$ s, in the presence of laminarin. For rCtD-Ole e 9, mean values are $\langle T_1 \rangle = 0.83$ s; $\langle T_2 \rangle = 0.091$ s and $\langle T_1 \rangle = 1.26$ s; $\langle T_2 \rangle = 0.061$ s, in the absence and presence of laminarin, respectively.

Hydrodymamic data of free rCtD-Fra e 9 and rCtD-Ole e 9 and their complexes with laminarin.

It is very well known that $\langle T_1/T_2 \rangle$ can be used to estimate the molecular correlation time (τ_c) in globular systems with isotropic tumbling. However, contributions of flexible tails (low NOE values) or regions affected by exchange processes can introduce distortions in the estimations and it is good practice to exclude these residues from the calculation process. In these cases, a good approach to obtain τ_c is the use of theoretical hydrodynamic calculations (Pérez-Cañadillas, Guenneugues et al. 2002), as described in the Materials and Methods section. The calculated correlation times were 6.5 ns for rCtD-Ole e 9 and 5.6 ns for rCtD-Fra e 9. In the complex with laminarin, values were 10.9 ns and 10.3 ns, respectively. The theoretical molecular masses of the complexes, on the basis of the τ_c value and extrapolating from known empiric values, were 16.9 kDa for rCtD-Fra e 9 and 17.9 kDa for rCtD-Ole e 9. These values agree, within the error of this approach, with a 1:1 complex.

The hydrodynamic behaviour of rCtD-Fra e 9 and rCtD-Ole e 9 in the presence of laminarin was also studied by analytical ultracentrifugation to determine the homogeneity of the protein solutions, the association state and the stoichiometry of the complex protein-oligosaccharide.

For both complexes, the sedimentation velocity profiles showed an apparent single boundary that could be described in terms of a single sedimenting species

(**Figure 5A**), with experimental sedimentation coefficients of 2.34 S and 2.15 S for rCtD-Fra e 9 and rCtD-Ole e 9 (the corresponding standard values, $s_{20,w}$, were 2.24 S and 2.07 S, respectively). Therefore, the absence any other sedimenting species indicates that rCtD-Fra e 9 and rCtD-Ole e 9 behave as homogeneous systems in the presence of laminarin at the experimental conditions (see Material and Methods).

On the other hand, the association processes can be studied by equilibrium sedimentation experiments when the mass of the complex significantly differs from the mass of the isolated components. This analysis yields the buoyant molecular weight of the sedimenting species (M_b). To convert these values to absolute molecular masses (the average molecular mass, $M_{w,app}$) the partial specific volume (v) of the particle is required ($M_b = M_{w,app}(1 - \rho v)$, where ρ is the density of the solution). The partial specific volume of a protein can be easily calculated from its residue composition (Perkins 1986), but this is not the case for the laminarin or for the protein-carbohydrate complex. As it has been described for glycoproteins (Ghirlando, Keown et al. 1995, Lewis and Junghans 2000), in these situations the stoichiometry of the complex can be readily obtained from the buoyant molecular masses instead of the $M_{w,app}$. Therefore, the theoretical buoyant masses of the rCtD-Fra e 9 and rCtD-Ole e 9 proteins were calculated and compared with the experimental values (**Table 1**).

The experimental equilibrium profiles were fitted to an ideal model (single species model) yielding very similar results at different rotor speeds (19,000, 22,400 and 33,000 rpm). The distribution of the residuals and the mass conservation for each velocity confirmed the homogeneity of the samples (**Figure 5B, 5C**), in agreement with the velocity results. The average value of buoyant mass from all experiments performed was of 5,901 ± 66.9 $g \cdot mol^{-1}$ for rCtD-Fra e 9 and 5,835 ± 153.4 $g \cdot mol^{-1}$ for rCtD-Ole e 9 laminarin mixtures, respectively (**Table 1**). These values are higher than those calculated for the free proteins alone and indicate that correspond to a protein-carbohydrate complex.

The molecular mass of glycoproteins has been proposed to be the sum of the molecular mass of the protein and the molecular mass of the carbohydrate portion ($M(1 - \varrho v) = M_p(1 - \varrho v) + M_c(1 - \varrho v)$ where p and c denote to the protein and the carbohydrate, respectively (Ghirlando, Keown et al. 1995, Lewis and Junghans 2000). Here this was assumed as valid for the formation of the protein-laminarin complex. Considering this, when the theoretical buoyant mass of the protein (M_b^1 in **Table 1**) is subtracted from the experimental buoyant mass (i.e., the buoyant mass of the complex;

M_b^2 in **Table 1**) results a resting buoyant masses of 2,677.8 g·mol^{-1} and 2,875.1 g·mol^{-1} should correspond to the molecular mass of the laminarin moiety. Taking into account that in the literature, the partial specific volume for carbohydrates is estimated to be in the range of 0.60-0.64 cm^3·g^{-1} (Durchschlag, 1986), we calculated a theoretical partial specific volume of ~1,936- 2,159 cm^3·g^{-1} for laminarin (Durchschlag 1986). Comparing this result with those obtained from the subtraction of M_b^1-M_b^2, and despite the significant experimental uncertainties, the results seem compatible with a 1:1 stoichiometry for the protein-laminarin complex, which is consistent with the NMR results.

DISCUSSION

Non-catalytic CBMs have been widely studied, not only because their fundamental importance as well as for their applications in many fields (Shoseyov, Shani et al. 2006). Although occasionally found as independent proteins (Hashimoto 2006), CBMs are usually part of large proteins and act to increase the activity of the linked catalytic modules by extending the substrate and putting it in close contact with the enzyme or *vice versa* (Boraston, Bolam et al. 2004). Despite recent advances, the molecular interactions that define the ligand specificity in CBMs and the mechanisms by which they recognize and select their substrates are not completely understood at the molecular level. Therefore we have studied and compared two similar CBMs, the C-terminal domains of two allergens with β−1,3-glucanase activity, rCtD-Ole e 9 and rCtD-Fra e 9. It is well known that NMR spectroscopy can provide unique information for weak protein–carbohydrate complexes in solution (μ-mM range) (Fernandez-Alonso, Diaz et al. 2012), such as the protein interaction surface and the groups involved in recognition (Garcia-Mayoral, Canales et al. 2013). In this study, NMR information, together with hydrodynamic data, have been useful to propose a model accounting for laminarin binding to aid in the understanding of the enzymatic mechanism of large β−1,3-glucanases.

In agreement with the sequence identity, the NMR results presented here have shown that free rCtD-Fra e 9 has the same secondary and tertiary structure as previously reported for free rCtD-Ole e 9 (Treviño, Palomares et al. 2008). Also the dynamic properties, from relaxation and ultracentrifugation experiments, have shown that both Ct-domains behave as monomers, and that the calculated correlation times (6.5 ns and

5.6 ns for Ole e 9 and Fra e 9 domains, respectively) are in agreement with other proteins of similar size and shape (García de la Torre, Huertas et al. 2000).

The structure of the proteins is maintained in the complexes. NMR data confirm the presence of well-folded domains upon binding and no dramatic chemical shift changes are concomitant with complex formation. This is in agreement with known examples of transient complexes with K_D values of the same order (mM) as the ones determined here (Garcia-Mayoral, Canales et al. 2013). rCt-Fra e 9 forms a lower affinity complex with laminarin than rCtD-Ole e 9 does. In fact, binding assays clearly shows that rCtD-Ole e 9 has higher affinity (K_D 0.032 mM) for laminarin that rCtD-Fra e 9 (K_D 1.1 mM). In this regard, the affinity of rCtD-Fra e 9 is closer to the one shown by Ole e 10 (K_D 4.9 mM) an independent CBM of the same family (CBM43) (Barral, Suarez et al. 2005), than to the homologous domain in olive tree pollen rCtD-Ole e 9.

These finding prompt the question of what are the structural bases for the different carbohydrate recognition properties of the homologous and phylogenetically related rCtD-Fra e 9 and rCtD-Ole e 9. To address this question, we have plot in the protein surfaces those residues that, according to NMR chemical shifts, are affected by laminarin binding (**Figure 4**). Interestingly, in both cases, all the perturbations lie in the same face of the molecule although the affected area is larger in the rCtD-Ole e 9 surface. This could mean that more residues are participating in binding or that the structural rearrangement after binding is larger in rCtD-Ole e 9.

We have looked in detail the residues affected by laminarin binding. As can be seen in **Figure 6**, the carbohydrate binding disturbs in both cases the inter-helical region of α1 and α2, although the number and residue type are significantly different in rCtD-Fra e 9 and rCtD-Ole e 9. It is very well known that the binding sites of the CBMs, which are generally composed of aromatic residues, are flat or platform-likes, and that the orientation of these residues is a key determinant of specificity (Boraston, Bolam et al. 2004). Thus it is not unexpected that the binding regions of rCtD-Fra e 9 and rCtD-Ole e 9 contain aromatic side chains. These are W372, H416, Y425, H427, W433 and F437 in rCtD-Ole e 9; and H420, Y423, W428, Y429 and F441 in rCtD-Fra e 9. Triptophan indole groups can interact strongly with sugars by forming stacking interactions (Asensio, Cañada et al. 1995). As we can see in **Figure 6**, in both molecules tryptophans are present in the interface. However, W428 in rCtD-Fra e 9 and W372 and W433 in rCtD-Ole e 9 are not equivalent, as they are placed in dissimilar environments and could interact with different parts of the sugar moiety. In addition, the presence of

charged residues (E409, R430 in rCtD-Ole e 9; and D393, R434 in rCtD-Fra e 9), that are oriented towards the binding place, can make contacts with the sugar moiety and determine the strength of the binding (**Figure 6)**. Also, both non-polar and polar interactions can contribute significantly to the complexation phenomena in both proteins, and the observed differences in the binding site can promote the stabilisation of different orientations of the sugar rings through the formation of different patterns of hydrogen bonds and hydrophobic interactions.

Little is known about the source of substrates, soluble or insoluble, that are naturally hydrolysed by these allergenic two-domain glucanases. Based on the affinity observed here, our results point to laminarin as being a preferred substrate, which would form a 1:1 complex with the CBM.

Furthermore, proteins that possess hydrolytic activity (e.g., cellulases and xylanases) encompass a complex molecular architecture comprising discrete modules (typically, a catalytic module and one or more CBMs), which are normally joined by relatively unstructured linker sequences (Shoseyov, Shani et al. 2006). This seems to be the case for Fra e 9 and Ole e 9. Although the 3D structure of the full proteins have not been elucidated, our structural characterization of isolated rCtD-Ole e 9 revealed a long unstructured N-terminal segment preceding its folded N-terminal domain (Treviño, Palomares et al. 2008). This unstructured segment corresponds to the linker connecting this CBM to the catalytic domain. From the mechanistic point of view, it has been proposed that the length and the flexibility of the linker play a critical role in the cooperative action of these domains driving the transition from an initial ‘‘open’’ configuration to a ‘‘closed’’ state. (Batista, Costa et al. 2011). In this regard, the observation that the recognition site is in one face of the protein structure should facilitate the accessibility of the substrate to the catalytic centre following the structural change proposed for the linker.

ACKNOWLEDGEMENTS

This work was financially supported with projects CTQ2011-22514 to M.B and SAF2011-26716 to R.R. and M.V. from the Ministerio de Economía y Competitividad and FCT-MEC, SFRH/BSAB/452/2004 to A.L.M., from the Fundação para a Ciência e a Tecnologia; and also by grant BES-2012-057717 to H.Z.

REFERENCES

Aramini, J. (2010). "http://www.nmr2.buffalo.edu/nesg.wiki/NMR_determined_Rotational_correlation_time."

Asensio, J. L., F. J. Cañada, M. Bruix, A. Rodriguez-Romero and J. Jimenez-Barbero (1995). "The interaction of hevein with N-acetylglucosamine-containing oligosaccharides. Solution structure of hevein complexed to chitobiose." Eur J Biochem **230**(2): 621-633.

Barral, P., C. Suarez, E. Batanero, C. Alfonso, D. Alche Jde, M. I. Rodríguez-García, M. Villalba, G. Rivas and R. Rodríguez (2005). "An olive pollen protein with allergenic activity, Ole e 10, defines a novel family of carbohydrate-binding modules and is potentially implicated in pollen germination." Biochem J **390**(Pt 1): 77-84.

Batista, P. R., M. G. Costa, P. G. Pascutti, P. M. Bisch and W. de Souza (2011). "High temperatures enhance cooperative motions between CBM and catalytic domains of a thermostable cellulase: mechanism insights from essential dynamics." Phys Chem Chem Phys **13**(30): 13709-13720.

Bolam, D. N., H. Xie, G. Pell, D. Hogg, G. Galbraith, B. Henrissat and H. J. Gilbert (2004). "X4 modules represent a new family of carbohydrate-binding modules that display novel properties." J Biol Chem **279**(22): 22953-22963.

Boraston, A. B., D. N. Bolam, H. J. Gilbert and G. J. Davies (2004). "Carbohydrate-binding modules: fine-tuning polysaccharide recognition." Biochem J **382**(Pt 3): 769-781.

Casacuberta, J. M., D. Raventos, P. Puigdomenech and B. San Segundo (1992). "Expression of the gene encoding the PR-like protein PRms in germinating maize embryos." Mol Gen Genet **234**(1): 97-104.

Castrillo, I., M. A. Treviño, O. Palomares, M. Rico, J. Santoro and M. Bruix (2006). "NMR assignment of the C-terminal domain of Ole e 9, a major allergen from the olive tree pollen." J Biomol NMR **36 Suppl 1**: 67.

Chen, C. and Z. Chen (2000). "Isolation and characterization of two pathogen- and salicylic acid-induced genes encoding WRKY DNA-binding proteins from tobacco." Plant Mol Biol **42**(2): 387-396.

Durchschlag, H. (1986). Specific volumes of biological macromolecules and some other molecules of biological interest. Thermodynamic data for biochemistry and biotechnology. E. Hienz H-J. Berlin, Springer**:** 45–128.

Farrow, N. A., R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay and L. E. Kay (1994). "Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation." Biochemistry **33**(19): 5984-6003.

Fernandez-Alonso, M. C., D. Diaz, M. A. Berbis, F. Marcelo, J. Canada and J. Jimenez-Barbero (2012). "Protein-carbohydrate interactions studied by NMR: from molecular recognition to drug design." Curr Protein Pept Sci **13**(8): 816-830.

García de la Torre, J., M. L. Huertas and B. Carrasco (2000). "HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations." J Magn Reson **147**(1): 138-146.

Garcia-Mayoral, M. F., A. Canales, D. Diaz, J. Lopez-Prados, M. Moussaoui, J. L. de Paz, J. Angulo, P. M. Nieto, J. Jimenez-Barbero, E. Boix and M. Bruix (2013). "Insights into the glycosaminoglycan-mediated cytotoxic mechanism of eosinophil cationic protein revealed by NMR." ACS Chem Biol **8**(1): 144-151.

Ghirlando, R., M. B. Keown, G. A. Mackay, M. S. Lewis, J. C. Unkeless and H. J. Gould (1995). "Stoichiometry and thermodynamics of the interaction between the Fc

fragment of human IgG1 and its low-affinity receptor Fc gamma RIII." Biochemistry **34**(41): 13320-13327.
Goddard, T. D. and D. G. Kneller (2005). "Sparky 3, University of California, San Francisco."
Guex, N., M. C. Peitsch and T. Schwede (2009). "Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective." Electrophoresis **30 Suppl 1**: S162-173.
Hashimoto, H. (2006). "Recent structural studies of carbohydrate-binding modules." Cell Mol Life Sci **63**(24): 2954-2967.
Huang, X. and W. Miller (1991). "A Time-Efficient, Linear-Space Local Similarity Algorithm." Adv Appl Math **12**: 337-357.
Huecas, S., M. Villalba and R. Rodriguez (2001). "Ole e 9, a major olive pollen allergen is a 1,3-beta-glucanase. Isolation, characterization, amino acid sequence, and tissue specificity." J Biol Chem **276**(30): 27959-27966.
Kay, L. E., D. A. Torchia and A. Bax (1989). "Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease." Biochemistry **28**(23): 8972-8979.
Laue, T. M., B. D. Shah, T. M. Ridgeway and S. L. Pelletier (1992). Analytical ultracentrifugation. . Biochemistry and Polymer Science. R. A. H. J. Harding SE, eds). Cambridge, UK., Royal Society of Chemistry, **:** 90–125.
Lewis, M. S. and R. P. Junghans (2000). Ultracentrifugal analysis of molecular mass of glycoproteins of unknown or ill-defined carbohydrate composition. Methods Enzymol **321:** 136–149.
Linthorst, H. J., L. S. Melchers, A. Mayer, J. S. van Roekel, B. J. Cornelissen and J. F. Bol (1990). "Analysis of gene families encoding acidic and basic beta-1,3-glucanases of tobacco." Proc Natl Acad Sci U S A **87**(22): 8756-8760.
Memelink, J., H. J. Linthorst, R. A. Schilperoort and J. H. Hoge (1990). "Tobacco genes encoding acidic and basic isoforms of pathogenesis-related proteins display different expression patterns." Plant Mol Biol **14**(2): 119-126.
Palomares, O., M. Villalba and R. Rodríguez (2003). "The C-terminal segment of the 1,3-beta-glucanase Ole e 9 from olive (Olea europaea) pollen is an independent domain with allergenic activity: expression in Pichia pastoris and characterization." Biochem J **369**(Pt 3): 593-601.
Pérez-Cañadillas, J. M., M. Guenneugues, R. Campos-Olivas, J. Santoro, A. Martínez del Pozo, J. G. Gavilanes, M. Rico and M. Bruix (2002). "Backbone dynamics of the cytotoxic ribonuclease alpha-sarcin by 15N NMR relaxation methods." J Biomol NMR **24**(4): 301-316.
Perkins, S. J. (1986). "Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences." Eur J Biochem **157**(1): 169-180.
Schuck, P. (2000). "Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling." Biophys J **78**(3): 1606-1619.
Shoseyov, O., Z. Shani and I. Levy (2006). "Carbohydrate binding modules: biochemical properties and novel applications." Microbiol Mol Biol Rev **70**(2): 283-295.
Torres, M. (2014). Taumatinas y b-1,3-glucanasas: dos familias de peoteínas de defensa de plantas implicadas en procesos alérgicos., Complutense University.
Treviño, M. A., M. F. García-Mayoral, P. Barral, M. Villalba, J. Santoro, M. Rico, R. Rodríguez and M. Bruix (2004). "NMR solution structure of Ole e 6, a major allergen from olive tree pollen." J Biol Chem **279**(37): 39035-39041.

Treviño, M. A., O. Palomares, I. Castrillo, M. Villalba, R. Rodríguez, M. Rico, J. Santoro and M. Bruix (2008). "Solution structure of the C-terminal domain of Ole e 9, a major allergen of olive pollen." Protein Sci **17**(2): 371-376.
Varea, J., J. L. Saiz, C. Lopez-Zumel, B. Monterroso, F. J. Medrano, J. L. Arrondo, I. Iloro, J. Laynez, J. L. Garcia and M. Menendez (2000). "Do sequence repeats play an equivalent role in the choline-binding module of pneumococcal LytA amidase?" J Biol Chem **275**(35): 26842-26855.
Verma, D. P. and Z. Hong (2005). "The ins and outs in membrane dynamics: tubulation and vesiculation." Trends Plant Sci **10**(4): 159-165.
Vogeli-Lange, R., C. Frundt, C. M. Hart, F. Nagy and F. Meins, Jr. (1994). "Developmental, hormonal, and pathogenesis-related regulation of the tobacco class I beta-1,3-glucanase B promoter." Plant Mol Biol **25**(2): 299-311.
Wishart, D. S., C. G. Bigam, A. Holm, R. S. Hodges and B. D. Sykes (1995). "(1)H, (13)C and (15)N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects." J Biomol NMR **5**(3): 332.
Wishart, D. S., B. D. Sykes and F. M. Richards (1991). "Relationship between nuclear magnetic resonance chemical shift and protein secondary structure." J Mol Biol **222**(2): 311-333.
Worrall, D., D. L. Hird, R. Hodge, W. Paul, J. Draper and R. Scott (1992). "Premature dissolution of the microsporocyte callose wall causes male sterility in transgenic tobacco." Plant Cell **4**(7): 759-771.

Table 1. *Characterization of the complex protein-laminarin by sedimentation equilibrium.*

Sample	MW* ($g \cdot mol^{-1}$)	v (20ºC) ($cm^3 \cdot g^{-1}$)	M_b^1 ($g \cdot mol^{-1}$)	M_b^2 ($g \cdot mol^{-1}$)	$M_b^1 - M_b^2$ ($g \cdot mol^{-1}$)
rCtD-Fra e 9	11,364.4	0.7055	3,223.2	5,901.0 ± 66.9	2,677.8
rCtD-Ole e 9	10,509.5	0.7074	2,960.6	5,835.7 ± 153.4	2,875.1

*Molecular mass from amino acid sequence.
[1]Theoretical value of buoyant molecular mass for rCtD-Fra e 9 and rCtD-Ole e 9; density of the solution: 1.0124 $g \cdot (cm^3)^{-1}$.
[2] Average of experimental buoyant masses of the proteins in the presence of 0.91 mM of laminarin. The error is the standard deviation of the mean, based on measurements at three rotor speeds (19,000, 22,400 and 33,000 rpm) at 25ºC

FIGURES

Figure 1. Structural data of rCtD-Fra e 9 and rCtD-Ole e 9. **A)** Sequence alignment. Identical residues are highlighted in green whereas similar residues are in orange. B) $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ conformational shift profiles for the CtD-Ole e 9 (top) (Treviño, Palomares et al. 2008) and rCtD-Fra e 9 (bottom). Regions with tendencies of secondary structure are colored cyan for β-strands and red for α-helices; whereas loops are indicated in black.

Figure 2. Carbohydrate binding activity of rCtD-Fra e 9 and rCtD-Ole e 9. A) C-terminal domains of Ole e 9 and Fra e 9 (2 µg) and BSA (0.7 µg) were electrophoresed under non-denaturing conditions in polyacrilamide gels in the presence (2.5 $mg \cdot mL^{-1}$) and absence of polysaccharides (-). B, Different concentrations of laminarin were assayed to determine the dissociation constant (K_D) of rCtD-Fra e 9 binding to laminarin, C. Empty dots correspond to rCtD-Fra e 9 and filled dots to rCtD-Ole e 9 data.

Figure 3. NMR titration data of rCtD-Fra e 9 and rCtD-Ole e 9 with laminarin. A) Superposition of the ^{1}H-^{15}N-HSQC spectra of rCtD-Fra e 9 acquired in the titration with laminarin. The spectra at different laminarin-rCtD-Fra e 9 molar ratios are represented in different colors (green = 0:1; blue = 1:1 and red = 2.5:1). Two examples of signal shift upon titration are shown on the right. B) Superposition of a region of the ^{1}H-^{15}N-HSQC initial and final spectra of laminarin-rCtD-Ole e 9 titration represented in different colors (blue = 0:1; red = 2.5:1). Signals that shift and one example of a signal that disappear (with a frame) upon titration are shown.

Figure 4. Protein-laminarin interaction surfaces. A) Averaged chemical shift variation ($\Delta\delta_{avg}$) of rCtD-Fra e 9 and rCtD-Ole e 9 upon the addition of an excess of laminarin. The most significant variations (more than the mean value) are highlighted in red, and signals that disappear following broading after binding are indicated with cyan line. Residues not assigned are represented with "*" and those that strongly overlap with "#". The elements of the secondary structure are shown on top. B) Representation of the chemical shift perturbation data in the protein surface of rCtD-Fra e 9 model. C) Representation of the chemical shift perturbation data in the protein surface of rCtD-Ole e 9 structure. A gradation of red intensities corresponds to the relative $\Delta\delta_{avg}$ perturbation values. Residues that disappear upon titration are shown in blue.

Figure 5. Velocity and equilibrium sedimentation experiments. A) Distribution of sedimentation coefficients for rCtD-Fra e 9 (black squares) and rCtD-Ole e 9 (grey circles) measured at 45,000 rpm, 25ºC. B) and C) Sedimentation equilibrium profile of rCtD-Fra e 9 and rCtD-Ole e 9 at 19,000 rpm, 25ºC. The continuous line is the fit of the experimental data to a single species model. Corresponding residual plots are below.

Figure 6. Model of rCtD-Ole e 9 and rCtD-Fra e 9 with the pentasaccharide laminaripentaose. A) The protein moieties of the docked rCtD-Fra e 9 and rCtD-Ole e 9 models are represented in ribbon and coloured in orange. The sugar is represented in sticks and coloured by a standard atom code. The orientation of the complexes was chosen to highlight the binding places. B) Stick representation of the aromatic and charged residues at the binding site in the ribbon model of the rCtD-Fra e 9 and rCtD-Ole e 9 backbones.

FIGURE 1

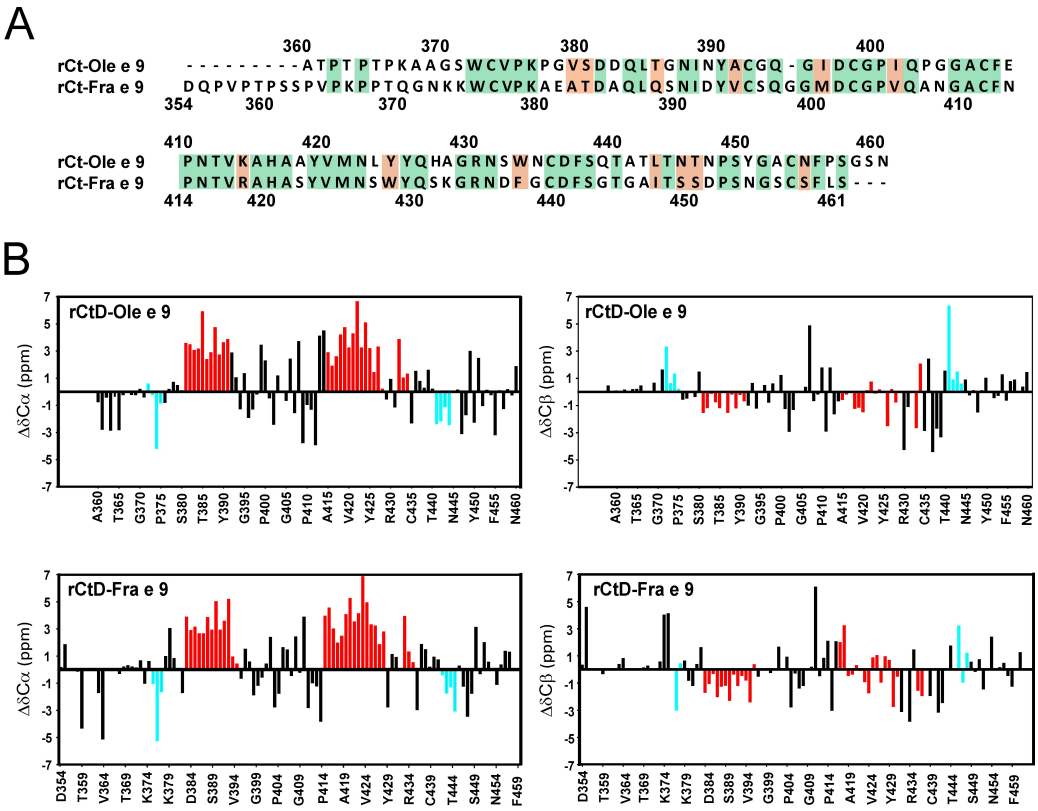

A
rCt-Ole e 9
rCt-Fra e 9
B
rCtD-Ole e 9
rCtD-Fra e 9
ΔδCα (ppm)
ΔδCβ (ppm)

FIGURE 2

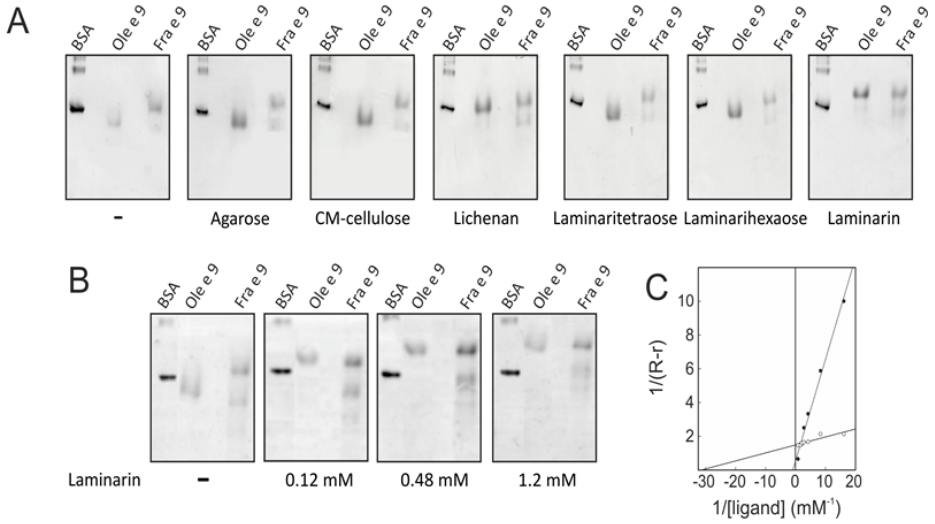

A
BSA
Ole e 9
Fra e 9
-
Agarose
CM-cellulose
Lichenan
Laminaritetraose
Laminarihexaose
Laminarin
B
Laminarin
0.12 mM
0.48 mM
1.2 mM
C
1/(R-r)
1/[ligand] (mM-1)

FIGURE 3

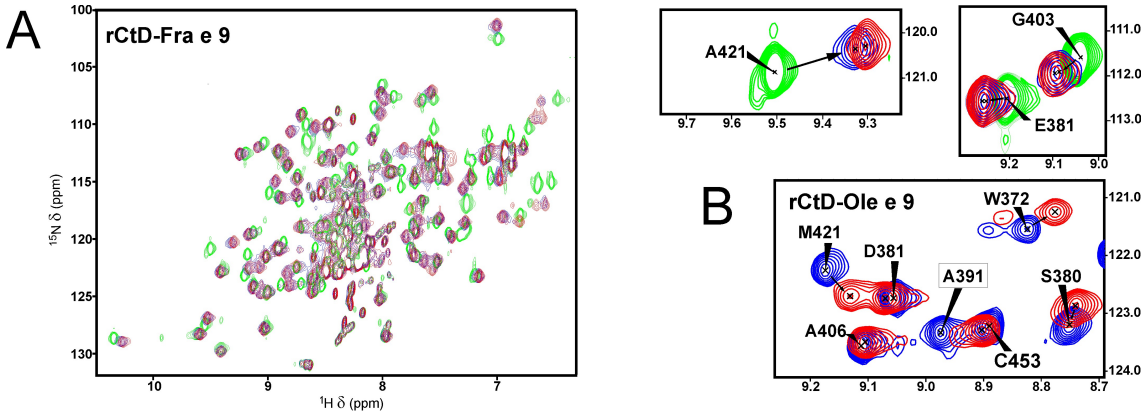

A
rCtD-Fra e 9
^{15}N δ (ppm)
^{1}H δ (ppm)
A421
G403
E381
B
rCtD-Ole e 9
W372
M421
D381
A391
S380
A406
C453

FIGURE 4

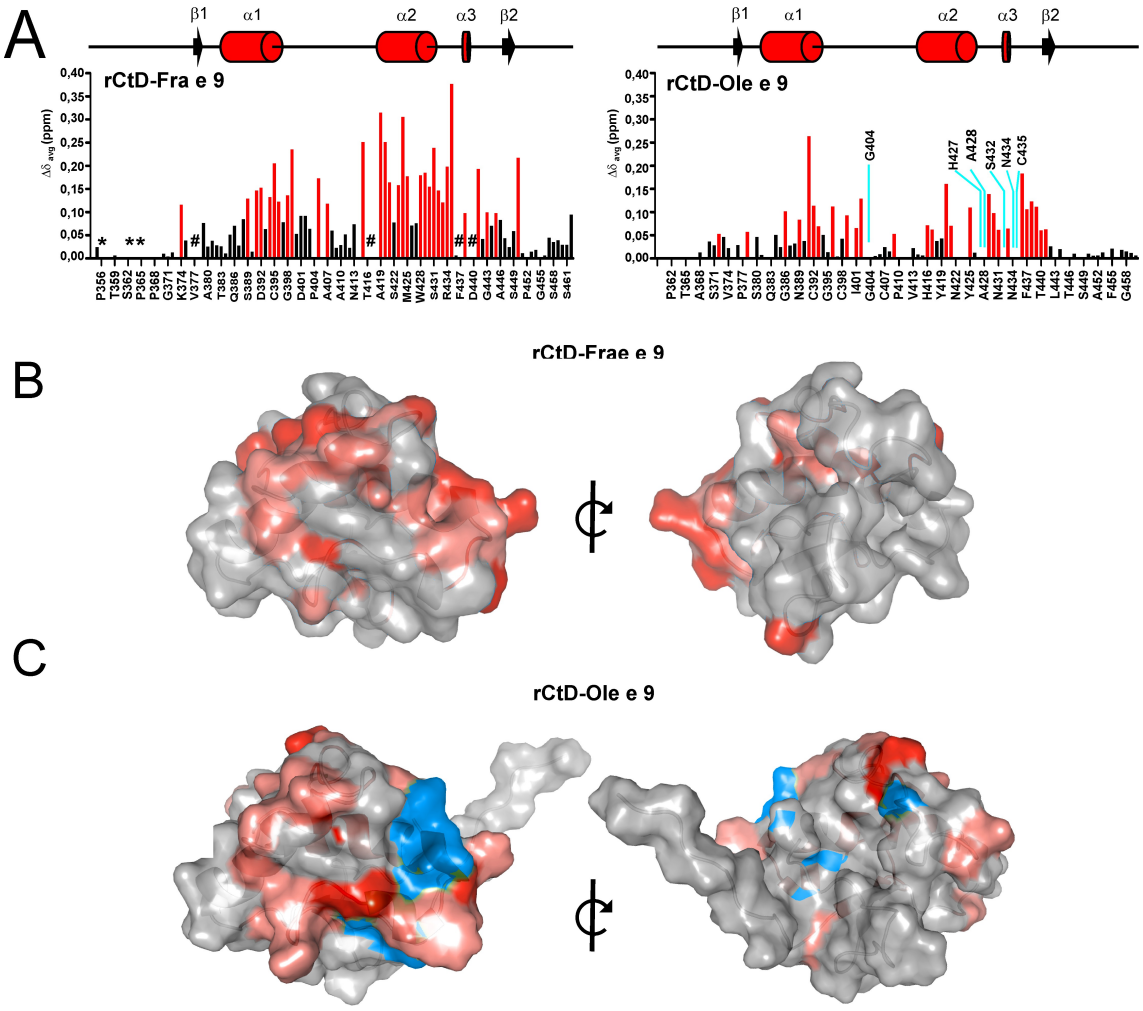

A
β1
α1
α2
α3
β2
rCtD-Fra e 9
rCtD-Ole e 9
$\Delta\delta_{avg}$ (ppm)
G404
H427
A428
S432
N434
C435
B
rCtD-Frae e 9
C
rCtD-Ole e 9

FIGURE 5

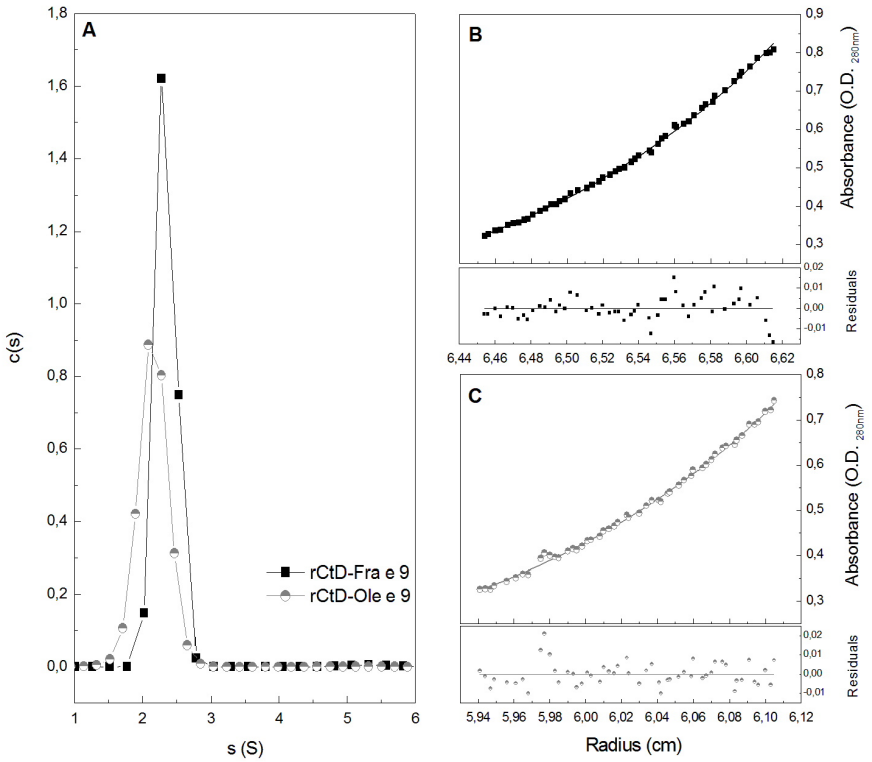
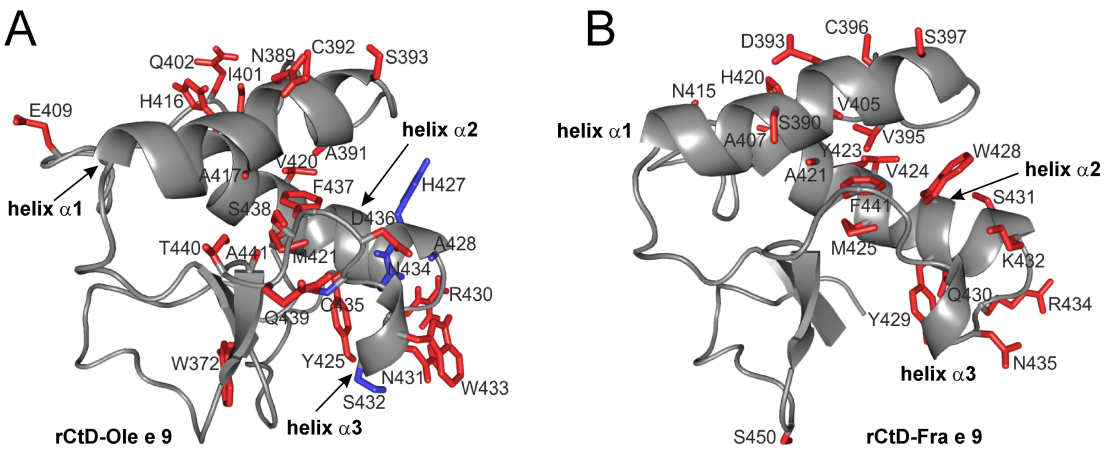

A
B
C
c(s)
s (S)
rCtD-Fra e 9
rCtD-Ole e 9
Absorbance (O.D. 280nm)
Residuals
Radius (cm)

FIGURE 6


A
Q402
N389
C392
S393
I401
H416
E409
helix α2
A391
V420
A417
F437
H427
helix α1
S438
D436
T440
A441
M421
A428
N434
R430
Q439
C435
W372
Y425
N431
W433
S432
helix α3
rCtD-Ole e 9
B
D393
C396
S397
N415
H420
V405
helix α1
S390
A407
V395
Y423
W428
A421
V424
helix α2
S431
F441
M425
K432
Q430
R434
Y429
N435
helix α3
S450
rCtD-Fra e 9